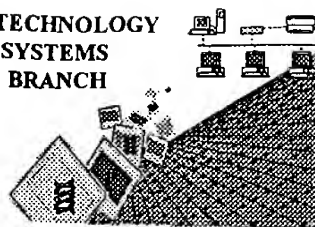


BEST AVAILABLE COPY

BIOTECHNOLOGY  
SYSTEMS  
BRANCH



M

**RAW SEQUENCE LISTING**  
**ERROR REPORT**

The Biotechnology Systems Branch of the Scientific and Technical Information Center (STIC) detected errors when processing the following computer readable form:

Application Serial Number: 09/869,891A  
Source: PL/09  
Date Processed by STIC: 6/21/2002

THE ATTACHED PRINTOUT EXPLAINS DETECTED ERRORS.

PLEASE FORWARD THIS INFORMATION TO THE APPLICANT BY EITHER:

- 1) INCLUDING A COPY OF THIS PRINTOUT IN YOUR NEXT COMMUNICATION TO THE APPLICANT, WITH A NOTICE TO COMPLY or,
- 2) TELEPHONING APPLICANT AND FAXING A COPY OF THIS PRINTOUT, WITH A NOTICE TO COMPLY

FOR CRF SUBMISSION QUESTIONS, PLEASE CONTACT MARK SPENCER, 703-308-4212.

FOR SEQUENCE RULES INTERPRETATION, PLEASE CONTACT ROBERT WAX, 703-308-4216.

PATENTIN 2.1 e-mail help: [patin21help@uspto.gov](mailto:patin21help@uspto.gov) or phone 703-306-4119 (R. Wax)

PATENTIN 3.0 e-mail help: [patin3help@uspto.gov](mailto:patin3help@uspto.gov) or phone 703-306-4119 (R. Wax)

TO REDUCE ERRORED SEQUENCE LISTINGS, PLEASE USE THE **CHECKER**  
**VERSION 3.1 PROGRAM**, ACCESSIBLE THROUGH THE U.S. PATENT AND  
TRADEMARK OFFICE WEBSITE. SEE BELOW FOR ADDRESS:

**<http://www.uspto.gov/web/offices/pac/checker>**

Applicants submitting genetic sequence information electronically on diskette or CD-Rom should be aware that there is a possibility that the disk/CD-Rom may have been affected by treatment given to all incoming mail.

Please consider using alternate methods of submission for the disk/CD-Rom or replacement disk/CD-Rom.

Any reply including a sequence listing in electronic form should NOT be sent to the 20231 zip code address for the United States Patent and Trademark Office, and instead should be sent via the following to the indicated addresses:

1. EFS-Bio (<<http://www.uspto.gov/ebc/efs/downloads/documents.htm>> , EFS Submission User Manual - ePAVE)
2. U.S. Postal Service: U.S. Patent and Trademark Office, Box Sequence, P.O. Box 2327, Arlington, VA 22202
3. Hand Carry directly to:  
U.S. Patent and Trademark Office, Technology Center 1600, Reception Area, 7<sup>th</sup> Floor, Examiner Name, Sequence Information, Crystal Mall One, 1911 South Clark Street, Arlington, VA 22202  
Or  
U.S. Patent and Trademark Office, Box Sequence, Customer Window, Lobby, Room 1B03, Crystal Plaza Two, 2011 South Clark Place, Arlington, VA 22202
4. Federal Express, United Parcel Service, or other delivery service to: U.S. Patent and Trademark Office, Box Sequence, Room 1B03-Mailroom, Crystal Plaza Two, 2011 South Clark Place, Arlington, VA 22202

Revised 01/29/2002

# Raw Sequence Listing Error Summary

## ERROR DETECTED

## SUGGESTED CORRECTION

SERIAL NUMBER: 09/869,891A

ATTN: NEW RULES CASES: PLEASE DISREGARD ENGLISH "ALPHA" HEADERS, WHICH WERE INSERTED BY PTO SOFTWARE

- 1      Wrapped Nucleics  
    Wrapped Aminos      The number/text at the end of each line "wrapped" down to the next line. This may occur if your file was retrieved in a word processor after creating it. Please adjust your right margin to .3; this will prevent "wrapping."
- 2      Invalid Line Length      The rules require that a line not exceed 72 characters in length. This includes white spaces.
- 3      Misaligned Amino  
    Numbering      The numbering under each 5<sup>th</sup> amino acid is misaligned. Do not use tab codes between numbers; use space characters, instead.
- 4      Non-ASCII      The submitted file was not saved in ASCII(DOS) text, as required by the Sequence Rules. Please ensure your subsequent submission is saved in ASCII text.
- 5      Variable Length      Sequence(s)      contain n's or Xaa's representing more than one residue. Per Sequence Rules, each n or Xaa can only represent a single residue. Please present the maximum number of each residue having variable length and indicate in the <220>-<223> section that some may be missing.
- 6      PatentIn 2.0  
    "bug"      A "bug" in PatentIn version 2.0 has caused the <220>-<223> section to be missing from amino acid sequences(s)     . Normally, PatentIn would automatically generate this section from the previously coded nucleic acid sequence. Please manually copy the relevant <220>-<223> section to the subsequent amino acid sequence. This applies to the mandatory <220>-<223> sections for Artificial or Unknown sequences.
- 7      Skipped Sequences  
    (OLD RULES)      Sequence(s)      missing. If intentional, please insert the following lines for each skipped sequence:  
    (2) INFORMATION FOR SEQ ID NO:X: (insert SEQ ID NO where "X" is shown)  
    (i) SEQUENCE CHARACTERISTICS: (Do not insert any subheadings under this heading)  
    (xi) SEQUENCE DESCRIPTION:SEQ ID NO:X: (insert SEQ ID NO where "X" is shown)  
    This sequence is intentionally skipped  
  
    Please also adjust the "(ii) NUMBER OF SEQUENCES:" response to include the skipped sequences.
- 8      Skipped Sequences  
    (NEW RULES)      Sequence(s)      missing. If intentional, please insert the following lines for each skipped sequence.  
    <210> sequence id number  
    <400> sequence id number  
    000
- 9      Use of n's or Xaa's  
    (NEW RULES)      Use of n's and/or Xaa's have been detected in the Sequence Listing.  
    Per 1.823 of Sequence Rules, use of <220>-<223> is MANDATORY if n's or Xaa's are present.  
    In <220> to <223> section, please explain location of n or Xaa; and which residue n or Xaa represents.
- 10      Invalid <213>  
    Response      Per 1.823 of Sequence Rules, the only valid <213> responses are: Unknown, Artificial Sequence, or scientific name (Genus/species). <220>-<223> section is required when <213> response is Unknown or is Artificial Sequence
- 11   J   Use of <220>      Sequence(s)   J   missing the <220> "Feature" and associated numeric identifiers and responses.  
    Use of <220> to <223> is MANDATORY if <213> "Organism" response is "Artificial Sequence" or "Unknown." Please explain source of genetic material in <220> to <223> section.  
    (See "Federal Register," 06/01/1998, Vol. 63, No. 104, pp. 29631-32) (Sec. 1.823 of Sequence Rules)
- 12      PatentIn 2.0  
    "bug"      Please do not use "Copy to Disk" function of PatentIn version 2.0. This causes a corrupted file, resulting in missing mandatory numeric identifiers and responses (as indicated on raw sequence listing). Instead, please use "File Manager" or any other manual means to copy file to floppy disk.
- 13      Misuse of n      n can only be used to represent a single nucleotide in a nucleic acid sequence. N is not used to represent any value not specifically a nucleotide.



PCT09

## RAW SEQUENCE LISTING

DATE: 06/21/2002

PATENT APPLICATION: US/09/869,891A

TIME: 11:27:38

Input Set : A:\PTO\_VSK.txt

Output Set: N:\CRF3\06212002\I869891A.raw

3 <110> APPLICANT: Universit<sup>①</sup> de Montr<sup>②</sup>al  
 4 Paquin, Bruno  
 5 Brukner, Ivan  
 6 Tremblay, Guy  
 8 <120> TITLE OF INVENTION: Process for the Generation of Oligonucleotide Libraries  
 9 (OLs) Representative of Genomes or Expressed mRNAs  
 10 (cDNAs) and Uses thereof  
 12 <130> FILE REFERENCE: 5593-00300  
 14 <140> CURRENT APPLICATION NUMBER: US 09/869,891A  
 15 <141> CURRENT FILING DATE: 2002-05-17  
 17 <150> PRIOR APPLICATION NUMBER: PCT/CA00/00047  
 18 <151> PRIOR FILING DATE: 2000-01-19  
 20 <150> PRIOR APPLICATION NUMBER: CA 2,259,745  
 21 <151> PRIOR FILING DATE: 1999-01-19  
 23 <160> NUMBER OF SEQ ID NOS: 5  
 25 <170> SOFTWARE: PatentIn Ver. 2.1  
 27 <210> SEQ ID NO: 1  
 28 <211> LENGTH: 67  
 29 <212> TYPE: DNA  
 30 <213> ORGANISM: Artificial Sequence  
 32 <220> FEATURE:  
 33 <221> NAME/KEY: misc\_feature  
 34 <222> LOCATION: (25)..(44)  
 35 <223> OTHER INFORMATION: n = any nucleotide a, c, g or t  
 37 <400> SEQUENCE: 1  
 38 gcctgttgatg agcctcctgt cgaannnnnnn nnnnnnnnnn nnnnttgagc gtttattctt 60  
 39 gtctccc 67  
 42 <210> SEQ ID NO: 2  
 43 <211> LENGTH: 24  
 44 <212> TYPE: DNA  
 45 <213> ORGANISM: Artificial Sequence  
 47 <220> FEATURE:  
 48 <223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic DNA  
 50 <400> SEQUENCE: 2  
 51 gcctgttgatg agcctcctgt cgaa 24  
 54 <210> SEQ ID NO: 3  
 55 <211> LENGTH: 23  
 56 <212> TYPE: DNA  
 57 <213> ORGANISM: Artificial Sequence  
 59 <220> FEATURE:  
 60 <223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic DNA  
 62 <400> SEQUENCE: 3  
 63 gggagacaag aataaacgct caa 23

**Does Not Comply**  
**Corrected Diskette Needed**

*do not  
 process  
 in the  
 CRT  
 software  
 program)*

*see item 11 on Error Summary Sheet*

## RAW SEQUENCE LISTING

DATE: 06/21/2002

PATENT APPLICATION: US/09/869,891A

TIME: 11:27:38

Input Set : A:\PTO\_VSK.txt

Output Set: N:\CRF3\06212002\I869891A.raw

66 <210> SEQ ID NO: 4  
67 <211> LENGTH: 24  
68 <212> TYPE: DNA  
69 <213> ORGANISM: Artificial Sequence  
71 <220> FEATURE:  
72 <223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic DNA  
74 <400> SEQUENCE: 4  
75 ttcgacagga ggctcacaac aggc 24  
78 <210> SEQ ID NO: 5  
79 <211> LENGTH: 23  
80 <212> TYPE: DNA  
81 <213> ORGANISM: Artificial Sequence  
83 <220> FEATURE:  
84 <223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic DNA  
86 <400> SEQUENCE: 5  
87 gggagacaag aataaacgct caa 23

RAW SEQUENCE LISTING ERROR SUMMARY  
PATENT APPLICATION: US/09/869,891A

DATE: 06/21/2002  
TIME: 11:27:39

Input Set : A:\PTO\_VSK.txt  
Output Set: N:\CRF3\06212002\I869891A.raw

Please Note:

Use of n and/or Xaa have been detected in the Sequence Listing. Please review the Sequence Listing to ensure that a corresponding explanation is presented in the <220> to <223> fields of each sequence which presents at least one n or Xaa.

Seq#:1; N Pos. 25,26,27,28,29,30,31,32,33,34,35,36,37,38,39,40,41,42,43,44